

This document is published in:

Casillas, J. et al. (eds.), 2012. *Management Intelligent Systems: First International Symposium* (Advances in Intelligent Systems and Computing, 171), Springer, pp. 47-56.

DOI: 10.1007/978-3-642-30864-2_5

© 2012 Springer-Verlag Berlin Heidelberg

ContextCare: Autonomous Video Surveillance System using multi-camera and smartphones

Gonzalo Blázquez Gil, Alvaro Luis Bustamante, Antonio Berlanga and José M. Molina
Group of Applied Artificial Intelligence, University Carlos III of Madrid,
Colmenarejo, Spain.
Email: {gbgil,aluis}@inf.uc3m.es, {aberlan,molina}@ia.uc3m.es

Abstract In the future, Ambient Intelligence (AmI) technology could assist people autonomously and interpret their intentions. Current technology can already be used to recognize the presence of a person in a private or public space and trigger an automatic response or reaction depending on the user activity. This work describes ContextCare, an extension for an video surveillance system in a health care scenarios based on activity recognition using sensor smartphones. Both systems are coordinated using ECA parading.

1 Introduction

According to the current financial crisis situation around the world, reduced budgets is a real challenge for governments. In this case, to freeze or reduce the healthcare budget is a priority while at the same time the service improves its quality. Thereby, AmI applications aim to contribute to reduce costs and offer a better and more efficient services. Anytime and anywhere assistance requires several underlying mechanisms and tools [8]: ranging from wireless-enabled monitoring, location systems that permit us to identify where they are located in case of needed.

There are many potential uses for AmI in Business management scenarios [4], however, ContextCare architecture is focused on human resources monitoring. Especially, hospitals and nursing homes scenarios where it possible to increase the efficiency of their services by monitoring patient's health, progress, and routines through the analysis of their activities, decreasing budgets.

Activity recognition is traditionally carried out through video systems like those described in [3] and [7]. However, video activity recognition systems present few problems like: huge computational cost, speedy bandwidth to transmit information, and complex techniques to interpret video data. Recent researches in activity recognition shows that Micro-Electro-Mechanical Systems (MEMS) is becoming another way to face with activity recognition problem [1, 5].

Normally, activity recognition researches based on MEMS are intrusive, for that reason, actual researches try to collect unobtrusively MEMS data using smartphones [6]. Smartphones may obtain and process physical phenomena from embedded sensors (accelerometer, gyroscope, compass, magnetometer, proximity sensor, light sensor, GPS, etc.) and transmit this information to remote locations without any human intervention.

Taking into account these advantages, it may be possible to consider a smartphone like a non-intrusive device to obtain people activities, since smartphones experience almost the same physical forces, temperature and noise that the person who carries them out. So, if you track mobile phone actions you are tracking person actions [11].

The huge potential of the smartphones devices has motivated us to design a framework that can intelligently capture different sensory data from a smartphone in real-time in order to coordinate a video surveillance system. ContextCare is capable to react and adapt video surveillance systems functionalities depending on the human resources context actions. Smartphone are not considered a mere communication device, otherwise it is introduced into an autonomous platform which allows to monitor efficiently patients with a video surveillance system.

ContextCare architecture rely on the ECA (Event-Condition-Action) paradigm which is composed by a defined set of reactive rules working over an event-driven architecture [10].

In this paper is presented the ContextCare architecture as an event-driven Semi-supervised video surveillance system that involves the use of smartphones and visual sensors.

2 Related works

ContextCare architecture is an monitoring system that englobes two different architectures: inContexto [6] which monitors user activity and location and video surveillance system [2] which uses inContexto information to focus on an anomalous situation reducing the time spend by security personnel in front of a screen.

2.1 InContexto: Architecture Definition

inContexto is a multimodal architecture which infers actions from users who carries out a smartphone (Figure 1).

A low-level sensing module continuously collects relevant information about the user activities. Thanks to Android OS that provides background processing is possible to run services without human control since the presented architecture is developed to obtain physical actions in a non-intrusive way. Data collection level gathers single raw data from smartphone sensors (Accelerometer, Gyroscope, GPS and so

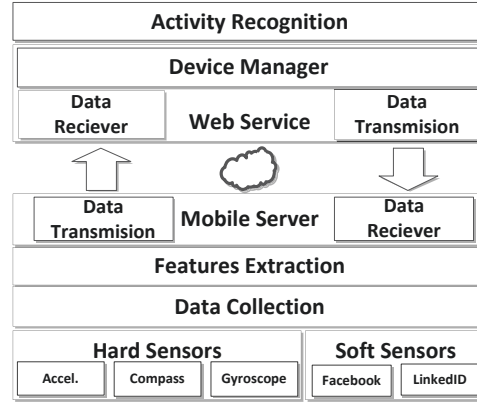


Fig. 1 inContexto architecture overview.

on) in order process and transform in features. Pre-processing is often required to improve performance, removing noise and redundancy in measurements.

Features extraction level which is also implemented in the smartphone, involves the extraction of symbolic features from smartphone sensor data. Hence, the objective of feature extraction is to represent an activity with the main characteristics of a data segment. In [5] has been defined the best features to infer user activity (Signal mean and variance, correlation between axes and signal energy).

Activity recognition module uses selected features to infer what activity is the user engaged in. In this component is implemented a J48 classification tree which was off-line trained.

2.2 Video Surveillance architecture

The Video Surveillance architecture is basically a system that allows the control of PTZ cameras by local or remote processes. In this case, this architecture mainly resides in the *Sensor Manager*, which is the responsible of attending control flows (allowing the positioning of PTZ cameras), and streaming video sequences to the different terminals. This component is used by ContextCare for monitoring patients, allowing an easy management of the underlying cameras.

The *Sensor Manager* used by ContextCare is the responsible of video acquisition, compression, and transmission, as well as to handle the communication protocols to perform the different movements in the cameras. The internal organization of this component is briefly outlined in figure 2.

There are two main functions that are controlled by the *Sensor Manager* for each video camera. The first one is related with the control of the camera and its move-

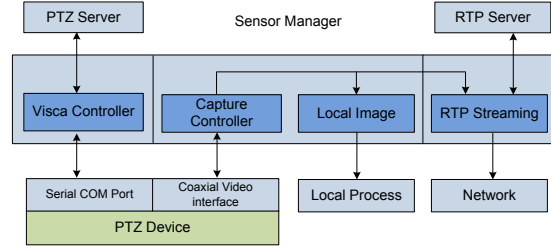


Fig. 2 Sensor manager for local/remote camera control

ments. As we can see in figure 2, there is a PTZ server which allows other processes to interact with the camera, so this controller can provide a standard interface to control homogeneously any underlying device. It has defined some high-level operating primitives like *goTo X Y*, *zoom amount*, etc. These high-level primitives are exposed as a non-connection oriented UDP Server, with a simple request-response protocol in the client-server computing mode.

The second main feature is about video acquisition and transmission. Generally, the access to limited resources like video devices is a problem if share the information between different terminals or systems is a requirement. In order to solve this issue we have defined two different strategies depending on the video destination. For local processes running on the same computer it is created a shared region of non-paged memory which can be accessed to retrieve the latest video frames. On the other hand, for remote processes, it is used a JPEG2000 [9] compression and real-time streaming system [2] with the aim of provide frames with the minimum delay. This allows to transmit real-time video sequences to remote processes like operators, agents, backup systems, mobile phones, etc.

3 Proposed Architecture: ContextCare

The main functionality of the ContextCare Surveillance Platform (See figure 3) is to connect seamlessly patients and the health professionals, reporting to the latter alerts and health measurements obtained from the patients. Besides, ContextCare is designed to manage video surveillance flow and determine which emergency member is the most suitable to look after the patient taking into account patient health context. Next subsection introduces the characteristics of the scenario where ContextCare architecture may be deployed and its architectural particularities.

3.1 ContextCare Scenario Model

Our interaction scenario considers a space populated with video cameras responsible of monitoring people. As we previously said, video surveillance systems are hardly tedious for security personnel due to within this scenario, there are plenty of people to be tracked.

The architecture is implemented following a event-driven model where patient context (event) triggers actions depending on a given condition (ECA rules). An ECA rules is divided in three different parts: the *event* is the signal that triggers a set of rules; the *condition* which if is satisfied makes the execution of the rule to continue and finally, the *action* defines the execution flow of a process.

ECA Rules are created by experts (security personnel or emergency services) and represent patient emergency situations. ContextCare architecture contains three different applications to deliver the global functionality:

- *Video surveillance system*: has been developed in C++ modules (as described in the previous works section) and it was evaluated in previous works [2]. Summarizing, the frame-rate obtained is the same as the video sensor provides (25 FPS).
- *Patient smartphone application*: The technological platform in the current prototype is an Android smartphone. Patient Smartphone app mainly consist in a monitoring service implementing inContexto architecture which collect patient sentinel data and establish user activity and location. Besides, is also able to reason with the raw sensor data to identify higher level information including patient activities or even though beat rate, temperature, etc.
- *Security mobile terminal (SMT)*: This application provides the necessary tools to access patients information. When a event triggers a rule, SMT receive a message which contains: patient profile, video tracking about patient situation, patient location and the event that was happened. Moreover, it is also provided, which allows the emergency services to evaluate the most recent medical details obtained from sensors, perform new measurements, and communicate with the caretakers.

Consider a patient in a nursing home wearing a smartphone with the normal sensors and blood pressure sensor connected. inContexto architecture captures every sensor data and generates patient context information. This information or event may be dispatched a rule (ECA rules) depending on the ECA rules condition. For example, an unusual blood pressure level triggers one rule and provokes a set of actions to be followed, sending a message to his surveillance system or even his family members, physician, emergency services, friends or colleagues.

3.2 Rule Manager Component

Rule Manager component is probably the most important component of the ContextCare architecture. it involves many IT disciplines like database access, knowl-

edge reasoning, video camera resource management and ECA rules evaluator. ECA model allows bidirectional communication between users (patients and security personnel) and video surveillance system in order to make possible to control and co-ordinate human resources and video cameras.

Rule Manager is a centralized server where the whole surveillance system is managed. Rule manager component's inputs are ECA rules provided by surveillance personnel which has the form: ON *event* IF *condition(s)* DO *action(s)*. Besides, user context activity information from inContexto architecture which follows this structure: (*userID*, *action*, *location*). On the contrary, rule manager outputs are guidelines (ECA rules actions) to carry out by the flow control component, coordinating cameras and also inform to the surveillance staff that something wrong is happening.

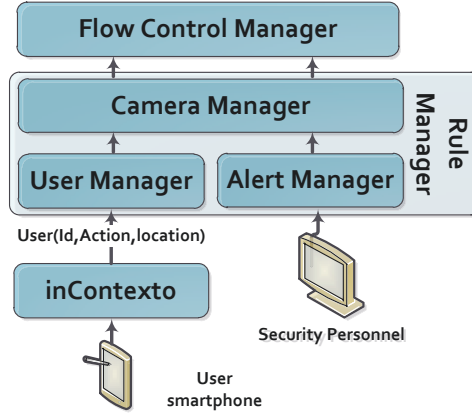


Fig. 3 ContextCare architecture: Rule Manager Component, video surveillance system and inContexto.

When ContextCare receives a new user context action from inContexto architecture, ECA rules is constantly evaluated in order to detect configured events, executing the associated actions if the conditions are fulfilled. In that case, first of all a message sequence is activated to notify from the users to Security personnel that something is happening. Later, ContextCare architecture decides which camera is the most propitious to track the situation. Finally, the system starts to monitoring autonomously the involved person.

3.2.1 User Manager

User Manager receives and stores every user context action and location generated by inContexto architecture. This component manages inContexto User State Mes-

sage (USM) and also stores into a database with the purpose of providing this information to camera manager which will check if any alert is active. inContexto USM is composed by the following attributes:

- *User id* which consist in a string which identifies every user, in this case smartphone direction IP. The Id permits to determines if the performed action is allowed to this person or not.
- *Action*, this field contains a concrete action or sensor value (depending on the embedded sensors on the system). Smartphone normally provides inertial sensors, however, inContexto allows to connect other sensors such as blood pressure, heart beat sensor, and so on.
- *Location* this field depicts the place where the user is in that moment. The suitability of each method depends on whether the location is outdoors or indoors and also the technology used. This field necessarily contains an absolute position like GPS coordinates. It could be filled by a symbolic position like corridor, room number, the nearest access point or wherever.

3.2.2 Alert Manager

Alert manager aims to communicate ContextCare architecture with human resources staff (Security or emergency personnel) terminals. The set of ECA rules are configured by an expert and define those events a flow control manager should be aware of.

As we explain previously, ECA Rules are composed by by three fields (events, conditions, actions) and they are described below:

- *Events*: describes a situation (user activity or location in this case) to which the rule may be able to respond. Events can be essentially divided into two categories: (i) primitive events which correspond to elementary occurrences and (ii) composite events that are composed for more than one primitive events.
- *Conditions*: specifies the conditions to trigger the ECA rule. Once the result of the condition evaluation is true, the condition is satisfied and the action field is executed.
- *Actions*: describes the task the rule considers relevant to the event and the condition. Actions field indicates the subsequent activities if the condition is satisfied.

Alert manager component generates two different responses, the first one is ECA rules created from the human resources terminal to Camera Manager and the second one is the alert protocol when an rules is triggered in the Camera Manager Component.

3.2.3 Camera Manager

Camera Manager component contains ECA Rules engine which is responsible of generate alerts and guidelines to manage the video surveillance system. These rules

involves to evaluate 'online' conditions, i.e., those which require to access an external resource (user context information, video camera control), therefore it is mandatory a real time response.

Mainly, there are two ways to control video flow. The first one is manually via the main terminal. Human Resources moves manually the cameras looking for situation of interest. The second way to control video flow is automatically via actions defined by ECA rules manager. Hence, it operates video cameras three different actors but each one play different roles:

- *Human Resources* monitoring the video streamed by different cameras and controlling their orientation manually.
- *Mobile human resources* walking around the monitored area. This person has a smartphone where receives alerts with security problems.
- *Users* are the principal actor. They also have a smartphone which is used to track their actions. When a non-usual action takes place the mobile phone will launch an alert to Rule manager.

Camera Manager gathers user context actions and ECA rules to build a unified view of the scene. Besides, it creates a goal which represent the overall objective of the video surveillance system. For example, a goal may be *track user which ID is 10001* and Camera Manager send user information (location and action) to the control flow manager which will be decided which camera is the most suitable to track this person.

Summarizing, Camera Manager decides according to the environment situation and the ECA rules which goal is a priority, sorting a list of goals actions for the Control flow Manager and creating the alert to the Security personnel.

4 ContextCare Architecture Evaluation: Study Case

ContextCare architecture has been created to monitor autonomously dependent people (Elderly people, children, etc.) and give health care support to hospital patients. Taking into account that this architecture is an autonomous monitor system, depending on the ECA Rules depicted by the human resources, it is mandatory a real-time performance to efficiently manage the information given by sensors.

In order to better assist the evaluation process the ContextCare system deployment will take place firstly in the University Carlos III of Madrid for pre-evaluation. The working scenario consists in a visual sensor network which contains six Pant-Tilt-Zoom (PTZ) control video devices around two rooms, three in the corridor and the other ones in the main room.

The scenario is a public place, so every person who enter to the track zone was under system control, however, there are just four people (playing user role) connected to ContextCare Application. In this cases, inContexto was configured to recognize five different actions (Lying, Standing, Walking, Running and Falling down).

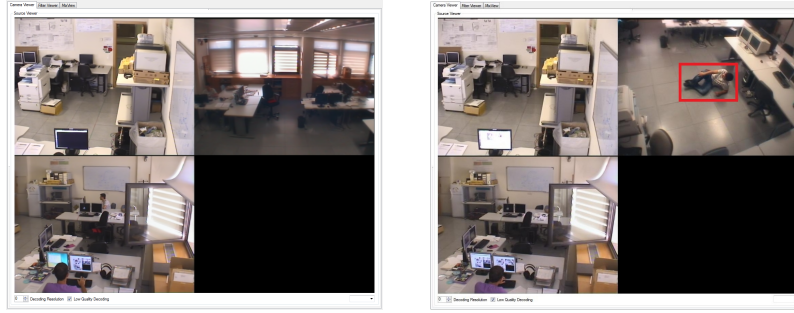


Fig. 4 ContextCare tracking a person who fell down on the floor. Figure a shows the monitored environment and Figure b shows a screen alert over the person.

In this case, an expert defines the ECA rules which describes what is considered a problem for each patient. Thereby, the ContextCare system may reach different rules like '*ON userContext:fall_down IF user:X DO follow*' or '*ON userContext:walking IF user=Y DO track*'). Every single action is detected by inContexto architecture and send it to ContextCare application where it is checked in the rule-based engine if it activates a rule.

First figure 4 depicts video surveillance system tracking. There are two persons sitting inside the room (bottom left camera), one of them stands up and start to walk. inContexto generates two different USM as follows:

- (*userID:10001,action:Stand,location:corridor*)
- (*userID:10001,action:Walking,location:(2,1)*)

First alert shows a symbolic location *corridor* and the second one an absolute position according to video surveillance coordinates. After that, the user fall down and the smartphone generates another USM which dispatches the next rule in the camera Manager Component. Alert manager rule engine checks the actual situation and the next ECA rule is dispatched.

- '*ON userContext:fall_down IF user:10001 DO follow AND Inform SP*'

Finally, alert system inform to camera manager component where is the user (Figure 4 shows the person on the floor inside a red rectangle) and also it creates the message to inform every security personnel.

5 Conclusions

AmI technology is developing fast and will promote a new generation on business management with some characteristics in the area of context awareness, anticipatory behavior and video surveillance. The presented architecture improves multicamera tracking applications performance, selecting the most suitable camera for any situation assessment during video tracking analysis. Using this new approach, the time

the security personnel spends in front of the screen is reduced, taking this time for other tasks.

ContexteCare architecture is well-suited in many human resources surveillance situations, for example the prevention of labor risks, management of human resources. However, we think that eHealth context is the most suitable situation.

Considered future works extending the development of the Activity Recognition system with more complex activities or even user emotional state, according to voice, face gestures, or typing patterns.

Acknowledgment

This work was supported in part by Projects CICYT TIN2011-28620-C02-01, CICYT TEC2011-28626-C02-02, CAM CONTEXTS (S2009/TIC-1485) and DPS2008-07029-C02-02

References

1. A. Avci, S. Bosch, M. Marin-Perianu, R. Marin-Perianu, and P. Havinga. Activity recognition using inertial sensing for healthcare, wellbeing and sports applications: A survey. In *Architecture of Computing Systems (ARCS), 2010 23rd International Conference on*, pages 1–10. VDE, 2010.
2. A. Bustamante, J. Molina, and M. Patricio. Multi-camera control and video transmission architecture for distributed systems. *User-Centric Technologies and Applications*, pages 37–45, 2011.
3. R. Cilla, M. Patricio, A. Berlanga, and J. Molina. A probabilistic, discriminative and distributed system for the recognition of human actions from multiple views. *Neurocomputing*, 2011.
4. D. Cook, J. Augusto, and V. Jakkula. Ambient intelligence: Technologies, applications, and opportunities. *Pervasive and Mobile Computing*, 5(4):277–298, 2009.
5. G. Gil, A. Berlanga, and J. Molina. Comparing features extraction techniques using j48 for activity recognition on mobile phones. *Distributed Computing and Artificial Intelligence: 9th International Conference*, 151:141, 2012.
6. G. Gil, A. Berlanga, and J. M. Molina. incontexto: A fusion architecture to obtain mobile context. In *Information Fusion (FUSION), 2011 Proceedings of the 14th International Conference on*, pages 1–8. IEEE, 2011.
7. J. Gómez-Romero, M. Serrano, M. Patricio, J. García, and J. Molina. Context-based scene recognition from visual data in smart homes: an information fusion approach. *Personal and Ubiquitous Computing*, pages 1–23, 2011.
8. G. López, V. Custodio, and J. Moreno. Lobin: E-textile and wireless-sensor-network-based platform for healthcare monitoring in future hospital environments. *Information Technology in Biomedicine, IEEE Transactions on*, 14(6):1446–1458, 2010.
9. A. Luis and M. Patricio. Scalable streaming of jpeg 2000 live video using rtp over udp. In *International Symposium on Distributed Computing and Artificial Intelligence 2008 (DCAI 2008)*, pages 574–581. Springer, 2009.
10. B. Michelson. Event-driven architecture overview. *Patricia Seybold Group*, 2006.
11. R. Want. You are your cell phone. *Pervasive Computing, IEEE*, 7(2):2–4, 2008.